Desktop Search Recommendation

People in today's business world are working with more and more electronic documents every day. This coupled with the ever increasing amount of business done via email, makes keeping everything organized in an "easy to find" manner almost impossible. There is a constant struggle to find needed information on demand, without losing productivity, while searching the different possible storage locations.

This problem is currently being addressed in two different ways, by two different types of software applications. The first, known as an "enterprise search tool" allows for searching across multiple different repositories such as intranet sites, document management systems, file servers, email mailboxes, etc. Many of these tools use a centralized index in order to ensure that data stored in centralized repositories is only indexed once, rather than each desktop creating its own index. Although these tools bring the search capability to many different repositories and attempt to create that "one stop search shop" for users, they also bring inherent costs along with them.

The second type of tool, known as a "desktop search tool" concentrates mainly on the searching of desktop and email content of an individual user, including the user's mapped network drives. These tools each build their own index which they run their searches against, with no ability to create a centralized index. Although more limited in functionality than the enterprise search tools, most desktop search tools are free of charge.

It is the study team's belief that the current demand within state government is for a free tool that can be used to enhance the searchability of user data. This coupled with Gartner's belief that there are three leading vendors in this arena; Google, Yahoo, and Microsoft, is the basis for the following recommendation, which concentrates on the free desktop search offerings from these three vendors. All three offerings were scored against a list of requirements that were developed by the study team and the results were recorded (Appendix A).

Google Desktop Search

The Google Desktop Search comes in two possible versions, the first being a standard version which is intended for home users. This version is very capable as far as its search capability. It allows for searching in email, calendar appointments, file system files (stored either locally or on a remote file server), etc. Its biggest drawback in a corporate environment is that it does not allow for centralized administration of its configuration.

The second version is known as the Enterprise Edition. This version has all of the same search capabilities as the standard one, with some additional options such as; integration with Google search appliances to allow for intranet searches, encryption of the index file, indexing of Lotus Notes content, and centralized administration of its configuration via Active Directory's Group Policy.

The centralized administration gives the ability to configure certain features and not allow the users to change these settings, ensuring that misconfiguration of a desktop search client doesn't cause issues related to security or performance. Although many options are available via the centralized administration the one major item missing is the ability to restrict indexing of network shares. Clients can be configured so that users can't add search locations, which would provide the same type of security as banning of network shares; however this would also hamper the user's ability to add local search locations.

Microsoft Windows Desktop Search

Microsoft's Windows Desktop Search is also a very capable desktop search tool. Its functionality compares very closely to that of the Google tool. It allows for the indexing of email, calendars, file system files, etc. It also allows for the searching of content inside email attachments, which the Google tool does not.

Centralized administration via Active Directory's Group Policy is available. This feature does not have quite the same granularity as the Google group policy template, however, it does allow for locking down a client in a way that would fit the current security and performance concerns of the State. It addresses the locking down of indexing locations slightly different than the Google tool in that it allows for the banning of network shares, while still allowing the users to add local index locations to their configuration.

Yahoo Desktop Search

The Yahoo Desktop Search tool is similar to the Google tool in that it comes in two possible versions. The first version is their regular Yahoo Desktop Search tool. This tool is limited in functionality much like the regular version of the Google tool. It does not allow for centralized administration and does not allow for the indexing of network shares. This missing functionality can be found in the other Yahoo version, known as X1. However this tool is not free of charge and therefore was not included in the scope of this study.

Recommendation

When implementing a desktop search tool caution must be taken in order to ensure that the deployment is done correctly, to prevent adverse effects on the State's environment. Although these tools seem harmless and can provide a source of increased productivity for users, the potential for violation of security policies and performance issues on the enterprise environment is great. Therefore the most important feature these tools must possess is centralized administration.

As important as we deem centralized administration, it is only useful if it is used to configure clients correctly. After doing extensive research and testing on the effects these tools can have on an environment, it was found that clients allowed to index against file servers and email servers can cause performance issues that should not be risked.

Because these tools are free to the public and can be downloaded and installed by anyone, IT staff must have control over their configuration. The centralized administration provided by Google and Microsoft allows for the desired configuration to be passed down via group policy and will not allow for users to make changes that are deemed as undesirable. Yahoo's lack of this functionality ultimately eliminates it from contention for this team's recommendation.

The distinctions between the Google and Microsoft tools are few. The biggest difference between the two is in one group policy setting. This setting identifies where a user is allowed to search. The Google tool does allow for locking down the search locations so that users can't add additional ones. This would allow IT staff to limit where their users can search remotely; however, it also prevents them from adding additional local search locations, which could inhibit the usefulness of the tool.

The Microsoft tool's group policy is not quite as granular as the one provided by Google, however it does allow for prohibiting the adding of remote search locations while still allowing users to add local search locations. This makes the tool useful at a local level, while still protecting the servers from adverse effects. The study team believes that this gives

Microsoft's tool an edge over the Google tool.

Therefore...it is the recommendation of this team to identify the "group policy aware" version of Microsoft Windows Desktop Search tool (version 2.06 or higher) as the standard desktop search tool. A group policy should be set as a domain policy (with a "no override" option set) and used to configure the clients so they can not index any network shares or email that resides on the State's Exchange servers. Indexing of local email stores such as POP3 mailboxes, .pst files and cached Exchange content (requires Outlook 2003) should be allowed. This configuration protects the file servers and email servers from any additional load brought on by indexing, while still providing as much local search capability as possible.

We would also recommend that all other desktop search tools be considered "banned" by the State and measures be put into place to ensure they are not installed on client workstations. Since an incorrectly configured client can easily cause issues and since these tools are easy to acquire and install, the only way to truly prevent these other tools from causing issues is to make sure they are not installed. One possible solution is to have the corporate anti-virus product recognize the other tools' files as "potentially unwanted programs" and not allow them to run.

The study team understands that this recommended configuration does inhibit the use of these tools for searching beyond the desktop. Even though some of them do provide for the ability to index and search remote locations, they do it in a manner which is inefficient and exposes the environment to possible performance issues. Therefore, if a need for search capability beyond the local desktop is desired, additional tools that are classified as enterprise search tools, should be investigated. These tools do a much better job of handling the searching of remote repositories; however they also have licensing charges associated with them.

SPECIAL NOTE:

It should be noted that care should be taken with the placement of the desktop search tool's indexes to ensure the necessary security of the information contained in the indexes is maintained. If a desktop tool is used to index information where access is controlled by a security mechanism different from the desktop's file security mechanism, information may be included in the indexes and may be unintentionally made available to entities not authorized to access the original information.

Appendix A

Requirement	Google Desktop Search	Microsoft Windows Desktop Search	Yahoo Desktop Search
Windows 2000/XP Compliant	Yes	Yes	Yes
Ability to configure and administer centrally (i.e. Group Policy)	Yes	Yes	No
Document security remains enforced	Yes	Yes	Yes
Use a minimum amount of client resources	Yes	Yes	Yes
Ability to search local drives	Yes	Yes	Yes
Ability to search network shares	Yes	Yes	No
Ability to search Outlook email (on Exchange server)	Yes	Yes	Yes
Ability to search Outlook Express	Yes	Yes	Yes
Ability to search Lotus Notes	Yes	No	No
Ability to search .pst files	Yes	Yes	Yes
Ability to search within attachments on emails	No	Yes	Yes
Ability to search Calendar and Contact information	Yes	Yes	Yes
Ability to search within attachments on Calendar entries	No	Yes	Yes
Ability to restricting of certain file types	Yes	Yes	Yes
Contain an embedded content viewer	Yes	Yes	Yes

Study Team Members

Mike Abel	Association of Counties
Wade Alvarez	Dept. of Transportation
Jerrold Arneson	State Court
Brian Bartz	Office of Management & Budget
Kory Hellman	Information Technology Dept.
Jayden Pascua	Land Dept.
Chris Sitter	Information Technology Dept.
Al Veit	Information Technology Dept.
Kevin Nosbusch	Information Technology Dept.